2–4 *MVA2000* IAPR Workshop on Machine Vision Applications, Nov. 28-30, 2000, The University of Tokyo, Japan

Finding Minimal Photo Hull for Image-based Rendering by Carving Space with Progressively Stricter Consistent Criterion

Chang-Ming Tsai [1,2] Chu-Song Chen [1]

[1] Institute of Information Science, Academia Sinica, Taipei, Taiwan
[2] Department of Computer Science and Information Engineering, National Taiwan University

Email: song@iis.sinica.edu.tw

Abstract

We propose the concept of the minimal photo hull which can produce the most consistent photographs among all of the photo hulls. In addition, it is also one of the optimal shapes making the minimal consistent level be the largest. Experimental results show that the minimal photo hull can be successfully applied for generation of convincing novel views.

I. Introduction

Given a set of images of an object taken from different viewing directions. Assume that (1) the images are taken with calibrated cameras whose poses and intrinsic parameters are known, and (2) the foreground (i.e., the silhouette of the object) and the background regions in each image have been segmented. In this paper, a 3D shape is *photo-consistent* if it satisfies both the image correspondence and the silhouette constraints. In detail, a color consistent criterion C(•) is defined to be a mapping which takes any *k* (*k>0*) colors as inputs and decide whether they are consistent or not. Let V be a 3D scene (or shape) defined by a finite, opaque, and possibly disconnected volume in space. Given a set of photographs. Let p be a point contained in V, and denoting U_P to be the set of pixels of the visible (i.e., un-occluded) projections of p on every photographs. The point p is photo-consistent under the consistent criterion C(•) if U_P contains no background pixels and the colors of all pixels in U_P are consistent under C(•). A shape V is photo-consistent if all of its points are photo-consistent and every foreground pixel is the projection of one of its point. The concept of photo-consistent shape was first proposed by Seitz and Dyer [5]. There may exist more than one photo-consistent shape with respect to a set of photographs.

The union of all photo-consistent shapes is also photo-consistent and is referred to as the *photo hull*. In the ideal case that no image noise is presented, the space-carving algorithm proposed by Kutulakos and Seitz [3] can ensure to obtain the photo hull. While image noise is considered, the space-carving algorithm may over-carve the shape, and thus fail to find a photo hull, if too strict a consistent criterion is used. In this paper, we propose a method which can ensure to find a photo consistent shape when image noise is taken into consideration. In particular, the photo-consistent shape obtained with our approach satisfies the *min/max optimization criterion*. That is, the reconstructed shape allows the maximal matching error of the color inconsistency to be the minimal.

II. Minimal Photo Hull

In the ideal case, it is intuitive that a set of colors is not consistent implies that each of its subsets is also not consistent (see Lemma 2 of [3]). In the case that image noise is considered, we also hope that the selected color consistent criterions satisfy the following monotonic property.

Definition 1 [Monotonic Consistent Criterion]: A consistent criterion C(•) is monotonic if a set of colors $G=\{g_1, g_2, \dots, g_k\}$ is consistent under C implies that any subsets of G are consistent under C.

For example, if the consistent criterion is selected that G is consistent iff $d(g_1, g_2, \dots, g_k) \le T$, where T is a threshold and $d(g_1, g_2, \dots, g_k)$ is the diameter of G [1], then it is monotonic. In particular, a consistent criterion C(•) is called *the loosest* in this paper if any sets of colors are consistent under C, and is *the strictest* if no sets of colors are consistent under C except that all colors in a set are exactly the same [2]. Conceptually, there is a mapping from the selected consistent criterion to the photo hulls. In addition, to deal with possible noise, it is better to give a measure of the tolerance levels to each consistent criterion. It inspired us to formulate a set of consistent criterions which are getting stricter as follows.

Definition 2 [Strictness of Consistent Criterion]: Given two color consistent criterions C_1(•) and C_2(•). C_1(•) is *stricter* than C_2(•) if a set of colors $\{g_1, g_2, \dots, g_k\}$ is consistent under C_1 implies that it is also consistent under C_2.

[1] The diameter of a set of points in the N-dimensional space is the largest Euclidean distance between each pair of the set of points.

[2] Here we assume that the surface reflection model is Lambertian. Nevertheless, our method can be extended for other surface reflection models.

Definition 3 [Progressively Stricter Consistent Criterions]: A class of consistent criterions $\mathbf{C}=\{C_h(\bullet) \mid 0\le h\le H\}$ (where H>0 is a constant) is *progressively stricter* if $0\le h_1\le h_2\le H$ implies that $C_{h2}(\bullet)$ is stricter than $C_{h1}(\bullet)$.

Let V be a shape in 3D. Given a class of consistent criterions $\mathbf{C}=\{C_h(\bullet) \mid 0\le h\le H\}$ which is progressively stricter. Then it is obvious that a point $p\in V$ is photo-consistent under C_t implies that it is photo-consistent under all of the consistent criterions less strict than C_t (i.e. than that contained in $\mathbf{C}_t = \{C_h(\bullet) \mid 0\le h\le t\}$). Denote Surf(V) to be the surface of V and let $Vis_V(p)$ be the collection of input photographs in which V does not occlude p. The visibility lemma [3] states that if $V'\subset V$ is a shape which also has p on its surface, then $Vis_V(p)\subseteq Vis_{V'}(p)$. By combining the above properties, a fundamental lemma about the non-photo-consistencies of points with respect to a set of progressively stricter consistent criterions is given below:

Lemma 1 [**Not Looser Photo-consistent Subset Lemma**] Let V be a shape in 3D. Given a class of consistent criterions $\mathbf{C}=\{C_h(\bullet) \mid 0\le h\le H\}$ which is progressively stricter and each $C_h(\bullet)$ is monotonic. A point $p\in surf(V)$ is not photo-consistent under C_t implies the following property: if $V'\in V$ is a photo-consistent shape under a consistent criterion which is not looser than C_t (i.e., which is contained in $\underline{\mathbf{C}}_t = \{C_h(\bullet) \mid t\le h\le H\}$), then $p\notin V'$. That is, every not looser photo-consistent subset of V do not contains p.

Intuitively, Lemma 1 suggests that a class of progressively stricter consistent criterions also exhibits a form of "monotonicity" which can be used for carving the space.

What we want to formulate is the mapping from a set of progressively stricter consistent criterions to the photo hulls. In particular, we hope to find the best one among these photo hulls – best in the sense that it is the most color-consistent with respect to the given photographs. To achieve this, we extend the concept of progressively stricter consistent criterions to as the follows:

Definition 4 [Completely Progressively Stricter]: A class of consistent criterions $\mathbf{C}=\{C_h(\bullet) \mid 0\le h\le H\}$ is *completely progressively stricter* if **C** is progressively stricter, $C_0(\bullet)$ is the loosest, and $C_H(\bullet)$ is the strictest.

When image noise is taken into consideration, there may exist no photo-consistent shapes under a given consistent criterion C if it is too strict. For example, in an extreme case, if the strictest consistent criterion is selected, then all voxels will be carved in the space-carving algorithm and the output will be an empty set. In the other extreme case, if the loosest consistent criterion is selected, then the output of the space-carving algorithm will be the visual hull [4] obtained via only the silhouette information. For those consistent criterions under which the photo-consistent shapes do exist, a unique photo hull can be obtained because it is defined as the union of all its photo-consistent shapes. In the following, we call a consistent criterion C *feasible* if there exists a photo hull with respect to C. Otherwise, C is *not feasible*.

Theorem 1 [Minimal Photo Hull Theorem]: Given a class of completely progressively stricter consistent criterions, $\mathbf{C}=\{C_h(\bullet) \mid 0\le h\le H\}$, and each $C_h(\bullet)$ is monotonic. Then there exists a discriminant value L that C_h is feasible for all $0\le h<L$, and is not feasible for all $L<h\le H$. In addition, the photo hull with respect to C_t is a subset of the photo hull with respect to C_h (defined as PH_h) for all $0\le h\le t<L$. In particular, the photo hull of the loosest criterion C_0, PH_0, is the visual hull.

We call the lower bound of the photo hulls, $\underline{PH}=\lim_{h\to L} PH_h$, the *minimal photo hull* in this paper. In particular, when no image noise is present, the minimal photo hull is equal to PH_H, the one with respect to the strictest criterion. The minimal photo hull is the photo hull computed under the strictest consistent criterion, and is a subset of every photo hulls with respect to the other feasible consistent criterions, $C_h(\bullet)$ $(0\le h<L)$. The stricter is the consistent criterion, the much more similar (to each other) are the colors of the set of pixels of the visible projections on the photographs. Therefore, the minimal photo hull can be treated as the best reconstruction among all feasible photo hulls because it produces the most consistent photographs to the given ones.

Definition 4 [Consistent Level]: Given a shape V, the *consistent level* of a point p contained in V, $CL_V(p)$, is defined to be the maximal value of h allowing p to be consistent under C_h.

It can be proven that the minimal photo hull also satisfies the following optimization property:

Theorem 2 [Theorem of Maximizing the Minimal Consistent Level]: The minimal photo hull is optimal in the sense that it is a shape allowing the minimal consistent level of the points contained in it to be the maximal.

III. Main Algorithm

Our purpose is to find the minimal photo hull with a set of photographs. By exploiting the monotonic property provided by the progressively stricter consistent criterions, we propose a systematic method which can find the minimal photo hull efficiently.

Algorithm [Space-carving with Progressively Stricter Consistent Criterion]:
Step 1: Initialize V to be the visual hull or a volume containing the true scene. Set $h \leftarrow \Delta h$.
Step 2: Let the consistent criterion be $C_h(\bullet)$. Set $V_{prev} \leftarrow V$.
Step 3: Find a voxel $v \in Surf(V)$ which is not photo-consistent. If no such a voxel can be found, then go to Step 4.
    3.1 Carve the space V by deleting the voxel v: $V \leftarrow V - \{v\}$.
    3.2 Go to Step 3.
Step 4: Project Surf(V) to every photographs. If there is any of the foreground pixel which is not projected by V, then stop and output $\underline{PH}^* = V_{prev}$.
Step 5: Set $h \leftarrow h+\Delta h$. Go to Step 2.

Basically, Algorithm 1 can find an approximation of the minimal photo hull. The accuracy of approximation depends on the incremental step Δh. A set of photo hulls, PH_0, $PH_{\Delta h}$, $PH_{2\Delta h}$ …, $PH_{\underline{L}}$ ($\underline{L}$ is an approximation of L) can be obtained under a getting stricter consistent criterions in each iteration of Algorithm 1. This algorithm can also be modified to obtain a more accurate estimation of the discriminant value H by reducing the step of by decreasing the search range Δh in a coarse-to-fine way.

In particular, Step 3 of Algorithm 1 is indeed the space-carving algorithm [3]. Hence, Algorithm 1 can be viewed as applying a getting stricter consistent criterion to the space-carving algorithm and verifies that whether every non-background pixel is the projection of a point in the reconstructed shape or not. Although the space-carving procedure is performed multiple times in Algorithm 1, a characteristic is that its computational complexity is the same with that of a single space-carving algorithm in terms of the number of voxels to be carved from the starting shape. It is because that every photo hull obtained under a looser consistent criterion is used as an initial volume to be carved for obtaining the photo hull under a stricter criterion.

A convenient way to realizing a set of progressively stricter consistent criterions in the implementation of Algorithm 1 is to thresholding an error function selected for deciding the color consistencies as described before. In addition, if Algorithm 1 is implemented in this way, we will show in the following that the minimal photo-hull also minimize the maximal error. Let $e(\bullet)$ be a mapping which takes any k (k>0) colors as inputs and generates a positive real number as an output. In particular, $e(\bullet)$ can be appropriately defined to measure the matching errors (or errors of inconsistencies) of a set of input colors, and thus is called a *matching error function* in our work. Given a threshold $T \geq 0$. We use the matching error function $e(\bullet)$ for the decision of color consistencies according to the following rule: A set of colors $G=\{g_1, g_2, \ldots, g_k\}$ is consistent if $e(g_1, g_2, \ldots, g_k) \leq T$, otherwise, it is not consistent. Hence, the smaller is T, the stricter is the consistent criterion [3]. In particular, the following property also holds for the minimal photo hull from the point of view of matching errors:

Lemma 2 [**Minimize the Maximal Matching Error**] The minimal photo hull is a shape allowing the maximal matching error of the points contained in it to be the minimal.

In practice, since the imprecision of the surface reflection model and the camera parameters will cause outliers of colors, a robust consistent criterion is better to be used. To allow the consistent criterions to be more robust, a solution is to select the second large or the third large distances, instead of using the largest distance (i.e. the diameter), between each pair of the colors contained in G. Such a selection can remove the affection of one or two outliers, and still allow the consistent criterion to be monotonic.

IV. Experimental Results

In this paper, the minimal photo hull is used for synthesizing novel views from a set of input images of an object taken from different viewing directions. In particular, to generate the novel-views with better quality, the technique of view-dependent texture mapping [2] is used.

Figure 1(a) shows 12 photographs of a Hello Kitty taken from different viewing directions. The box with some specially-designed feature lines and points under the Hello Kitty was used for camera parameter estimation. In our work, a cuboid-based calibration method [1] is used for camera parameter estimation (including both the intrinsic and the extrinsic parameters). The images are then segmented for extracting the silhouettes using the chroma-keying technique. Figure 1(b) shows some generated novel views (in fact, a novel video was generated) from the minimal photo hull with respect to the photographs shown in Figure 1(a). One can see that although the minimal photo hull is not equal to the real shape of the object, the generated novel views are still quite convincing.

---

[3] For example, if $e(g_1, g_2, \ldots, g_k)$ is selected to be the diameter of G, a consistent criterion $C_h(\bullet)$ ($0 \leq h < H$) can be defined as the one with $T= M(H-h)/H$ (where M is the largest possible value of the distance between two colors in the color space, e.g., $M=255\sqrt{3}$ for a standard RGB space).

V. Conclusions and Discussions

In this paper, we propose a new formulation in the photo-consistent shapes, the minimal photo hull. The minimal photo hull is the best among all feasible photo hulls in the sense that it produces the most consistent photographs to the given ones, and thus is very suitable for novel-view generation. In addition, it is also the optimal shape making the minimal consistent level be the largest (or equivalently, making the maximal matching error be the smallest). To find the minimal photo hull, an efficient algorithm is proposed in this paper by carving the 3D space with a set of progressively-stricter color consistent criterions. Our approach for 3D reconstruction makes full use of both silhouette and image-correspondence information contained in the given photographs.

The approach proposed in this paper is the first one which can find the photo-consistent 3D structure *in an optimal sense*. Experimental results show that our approach can be successfully applied for image-based rendering by generating very convincing novel views.

Acknowledgment

This work was supported in part by the National Science Council, Republic of China under Grant NSC-89-2218-E-001-004.

References

[1] C. S. Chen, C. K. Yu, and Y. P. Hung, "New Calibration-free Approach for Augmented Reality Based on Parameterized Cubed Structure," *Proceedings of International Conference on Computer Vision, ICCV'99*, Corfu, Greece, pp. 30-37, 1999.

[2] P. E. Debevec, C. Taylor, and J. Malik, "Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-based approach," *Proc. SIGGRAPH'96*, pp. 11-20, 1996.

[3] K. Kutulakos and S. M. Seitz, "A Theory of Space Carving," *Proceedings of International Conference on Computer Vision, ICCV'99*, 1999.

[4] A. Laurentini, "The Visual Hull Concept for Silhouette-based Image Understanding," *IEEE. Transactions on Pattern Analysis and Machine Intelligences,* pp. 150-162, 1994.

[5] S. M. Seitz and C. R. Dyer, "Photorealistic Scene Rendering by Voxel Coloring," *International Journal of Computer Vision*, 35(2), pp. 151-173, 1999.

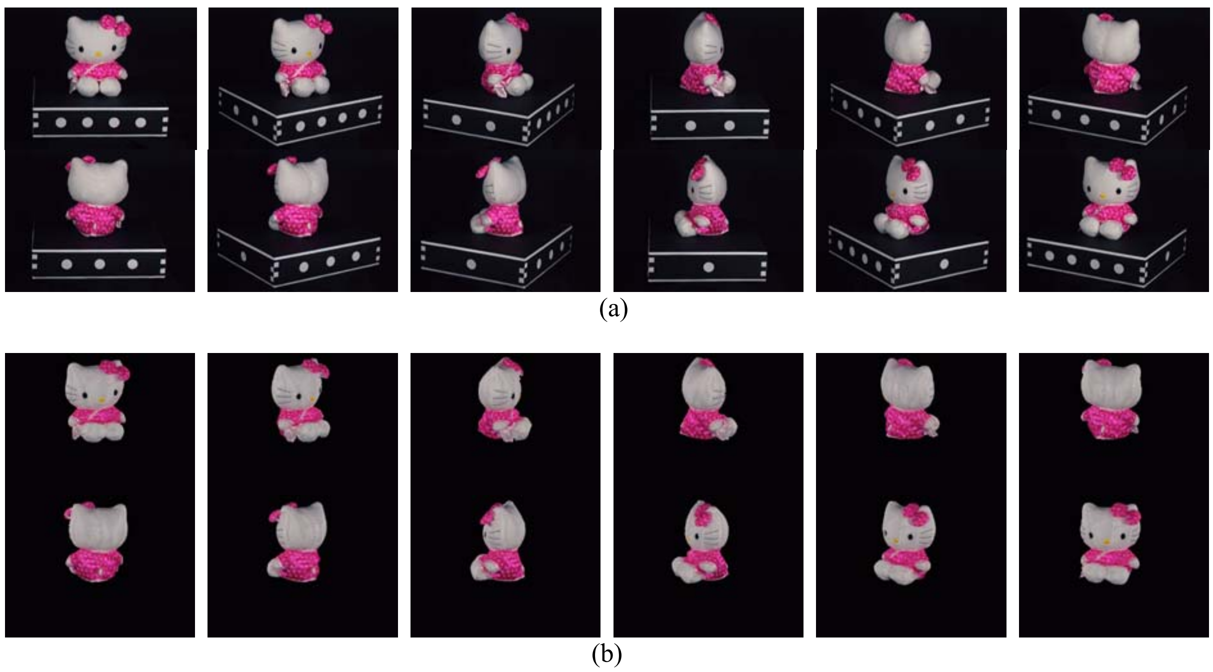

Figure 1. (a) Twelve photographs of a Hello Kitty taken from different view points. The box with some specially-designed feature lines and points under the Hello Kitty was used for camera calibration. The images are then segmented for extracting the silhouettes using the chroma-keying technique. (b) Some generated novel views (in fact, a novel video was generated) of the Hello Kitty via the minimal photo hull computed from the photographs shown in Figure 1(a).